On Data Augmentation for GAN Training

Year: 2020 | Citations: 309 | Authors: Ngoc-Trung Tran, Viet-Hung Tran, Ngoc-Bao Nguyen, Trung-Kien Nguyen, Ngai-Man Che

---

Abstract

Recent successes in Generative Adversarial Networks (GAN) have affirmed the importance of using more data in GAN training. Yet it is expensive to collect data in many domains such as medical applications. Data Augmentation (DA) has been applied in these applications. In this work, we first argue that the classical DA approach could mislead the generator to learn the distribution of the augmented data, which could be different from that of the original data. We then propose a principled framework, termed Data Augmentation Optimized for GAN (DAG), to enable the use of augmented data in GAN training to improve the learning of the original distribution. We provide theoretical analysis to show that using our proposed DAG aligns with the original GAN in minimizing the Jensen–Shannon (JS) divergence between the original distribution and model distribution. Importantly, the proposed DAG effectively leverages the augmented data to improve the learning of discriminator and generator. We conduct experiments to apply DAG to different GAN models: unconditional GAN, conditional GAN, self-supervised GAN and CycleGAN using datasets of natural images and medical images. The results show that DAG achieves consistent and considerable improvements across these models. Furthermore, when DAG is used in some GAN models, the system establishes state-of-the-art Fréchet Inception Distance (FID) scores. Our code is available (https://github.com/tntrung/dag-gans).